Review of “Essays on Early Tracking School System,” the dissertation research of Miroslava_Federičová

This dissertation investigates various effects of a school system with early tracking on achievement and school choices. The three essays consider gender differences in the response to tracking, the effects of entry examinations on effort and achievement, and the peer effects induced by tracking on students who choose not to apply to selective schools.

I find this dissertation to be thoughtful, careful, and compelling and congratulate Miroslava on the completion of a very good dissertation. Each chapter presents a thoughtful conceptual framework and develops appropriate methods to conduct the empirical analysis. The systematic assembly of empirical evidence provides comprehensive and convincing tests of hypotheses. This characteristic of the research is quite impressive across the dissertation.

In general the exposition is well-written. I suggest changing the title as it does not reflect the quality of writing in the dissertation. Perhaps “Essays on the Effects of Early School Tracking” or “Essays on Early school Tracking.” If the current title is preferred it the word “an” should be inserted before Early.

I now provide some comments on each of the papers. As I believe the dissertation passes the threshold I am not requesting that these be completed in order to pass. Rather I provide the comments for future revisions of the papers.

1. Gender Gap in Application to Selective Schools: Are Grades a Good Signal?

Be careful not to take evidence as fact in your work. The final paragraph on page 6 provides an example. Perhaps the second sentence could read “Evidence presented in Charness and Gneezy (2012) shows that girls tend to be more risk averse, and the findings in Spencer, et al (1999) suggest that girls face higher testing anxiety due to stereotype threat.

The signaling model raises some questions. First, measurement error typically affects variances and not simply means. Grade could be a noisy signal with classical measurement error for which the signal to noise ratio is lower for boys. In this case the more boys would be in the upper tale (and lower tale) of the grade distribution. Moreover, the coefficient for boys would be smaller in equation (1).

However, the model in this case does not consider classical measurement error, and I find it difficult to follow. There is a gender specific term and a general term. Clearly they are not centered around zero.

I suggest an alternative such as the following might be easier to follow. There are many variations, but it would seem to highlight the key ideas.

Admission test score is a function of only cognitive skill (cs). this is a simplification that could be modified.

Grade is a function of cognitive skill and other skills (n) in the manner described below.

Students predict admission score based strictly on grade.

The distribution of c is the same for boys and girls but the distribution of n is higher for girls than for boys. For a simple example consider 2 trivariate distributions for n and cs in which the variables take values of 1, 2, or 3 (1 is highest skill level and 3 is lowest). Thus there is the following 3 by 3 set of possible combinations of n and cs and the letter grades (American system where is best and D worst) produced by each combination:

		cs		
		1	2	3
	1	A	A	B
N	2	A	B	C
	3	B	C	D

Let the distribution of cs be the same for boys and girls and equal to (25%, 50%, 25%).

In contrast let the distribution of n be different for boys than for girls. Let it be (50%, 30%,20%) for girls and (30%, 40%, 30%) for boys.

To simplify let the distributions of c and n be independent (not realistic, as n affects knowledge acquisition and thus cognitive skill).

The probability of receiving an A would equal .25*.5 + .5*.5 + .25*.3 = 45% for girls, but .25*.3 + .5*.3 + .25*.4 = 32.5% for boys

The paper presents a model for individuals, but it is clear from this example that not all boys have the same signaling error and not all girls have the same error. On average boys will tend to understate their cognitive skill using grade as a signal relative to girls, but the relationship may be nonlinear depending upon the underlying distributions of n and c.

Given that you discuss other skills in the subsequent section, it seems appropriate to include a second skill in the theoretical framework rather than simply an error that is gender specific.

On page 19 in the top sentence, rather than saying that these are specific by gender you might say the distribution of these for girls is to the right of that for boys.

On page 20 be careful in the discussion of measurement error. IV works when the errors in the math and science tests are uncorrelated so NOT exposed to the same measurement error.

Is the probability of admission for applicants in the right side of figure 1.3 higher for boys? The boys distribution seems to be to the right of the girls. Is the difference significant? You might discuss this diagram.

In figure 1.6 the title to the figures should be predicted admission probability deciles

The paper makes strong statements about grades being a signal that skews beliefs. I do not believe there is any survey data to support this. Therefore, it is better to report that the "results are consistent with" or "the results suggest"

Figure 1.8 does not provide information on gender differences. It might be worth referencing this in terms of the differences in the grade distribution by probability of admission.

The phrase ineffective allocation of students on page 32 is awkward. Perhaps it is an inefficient allocation of students among schools assuming that admission probability is a good measure of the expected return to selective schooling for boys and girls.

Grammatical suggestions

1. Use student achievement or preferably just achievement. The plural achievements with the possessive students' is not generally used.
2. Top paragraph of page 10, omit "i.e." and second "decision" in first sentence
3. First rather than firstly under 1.6 results

The Impact of High-Stakes School-Admission Exams on Study Effort and Achievements: Quasi-experimental Evidence from Slovakia

This paper provides a lot of information about schools and students.

Rather than simply relying on dif-in-dif the analysis investigates the relationship between excess demand for selective school and test score. For completeness you could interact excess demand with DD and DY separately.

It does appear that class size fell in Slovakia relative to CR over time. This would tend to bias downward the effect.

I think the findings might generalize to high stakes tests of all kinds and not just tests to selective schools. An important issue is the longer-term effect and interaction with underlying skills. Would students learn more if there were high-stakes tests every year? Would students

not at the skill level for entrance to selective schools also respond to a test score incentive relevant for them? This paper probably cannot address these questions.

Children Left Behind: Self-confidence of Pupils in Competitive Early Tracking Environments

This is the least developed and compelling of the papers. Nonetheless, it is quite interesting and relevant for policy.

Be careful here with the discussion of the evidence. The statement that gender differences in math contribute to girls' underperformance seems too strong. Again, it might be better to say that evidence suggests or shows….

Interpretation of the measure as one of self confidence as opposed to self assessment seems questionable.

Should students be affected more by same-gender applicants?

Paper would benefit from a more comprehensive structure. Self-confidence is clearly endogenous, and yet that endogeneity is not addressed.

What is the role of the teacher in both amount learned and attitudes?

What is role of the family in the decision to apply? How does that interact with number of applicants?

Structure of self confidence variable is questionable.

Shouldn't applicant variable be measured as proportion applying in grade?

GPA is an outcome and interrelated with self confidence. The use of endogenous variables raises questions and merits additional discussion.